



研究与开发

一种自由视点视频系统中立体图像质量的评价方法

王士培, 彭宗举, 陈芬, 蒋刚毅, 郁梅

(宁波大学信息科学与工程学院, 浙江 宁波 315211)

摘要: 提出了一种自由视点视频系统中立体图像的质量评价方法。分别对参考图像与失真图像进行小波分解后用对比敏感度函数加权, 计算两者之间的平均结构相似度得到基于内容的质量分数; 在双目匹配区域中找到中心凹区域, 在该区域上计算图像之间的平均相位相似度, 得到基于深度感知的质量分数; 利用阈值提取匹配区域中为人眼所感知的失真区域, 与利用视差映射过来的另一视点图像的相对应区域计算平均结构相似度得到基于双目特性的质量分数。融合各部分质量分数得到最终图像质量。实验结果表明, 本算法评价结果与主观感知一致性较好, 更符合人眼视觉特性。

关键词: 自由视点视频; 立体图像; 虚拟视点; 人眼视觉特性; 深度感知

中图分类号: TP391.41

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2018038

A quality assessment metric for stereoscopic image in free viewpoint video system

WANG Shippei, PENG Zongju, CHEN Fen, JIANG Gangyi, YU Mei

Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China

Abstract: A quality assessment metric of stereoscopic image in free viewpoint video system was proposed. After the reference and distorted images being decomposed by wavelet then weighted by contrast sensitivity function respectively, the content score was obtained by calculating the average structural similarity. After the center region being found in the binocular matching region, the average phase similarity was calculated on the region to obtain depth perception. The distortion region perceived by the human eye in the matching region by using the threshold, the average structural similarity was calculated between this region and the corresponding region of the other view image by disparity mapping, then the binocular characteristic score was obtained. The final image quality can be obtained by fusing score of each part. The experimental results show that the proposed algorithm has better consistency with subjective perception and it is more in line with human visual characteristics.

Key words: free viewpoint video, stereo image, virtual viewpoint, human visual characteristic, depth perception

收稿日期: 2017-08-14; 修回日期: 2017-10-26

基金项目: 国家自然科学基金资助项目(No.61771269, No.61620106012, No.U1301257); 浙江省自然科学基金资助项目(No.LY16F010002, No.LY15F010005, No.LY17F010005)

Foundation Items: The National Natural Science Foundation of China (No.61771269, No.61620106012, No.U1301257), The Natural Science Foundation of Zhejiang Province of China (No.LY16F010002, No.LY17F010005, No.LY17F010005)

1 引言

4G/5G^[1-3]时代的来临,提升了用户的观影感受,也推进了电影行业的发展,尤其是近年来兴起的立体电影行业,观众通过佩戴立体眼镜来获得逼真的立体感,但受视点数量限制,不能任意选择视角进行观看,使观影具有了一定的局限性,因此,自由视点视频(free viewpoint video, FVV)^[4]系统应运而生。多视点加深度(multi-view video plus depth, MVD)是一种主流的三维场景表示方式,通过基于深度的虚拟视点绘制(depth image based rendering, DIBR)技术^[5]绘制虚拟视点,让用户自由地选择视角观看三维场景。现有的虚拟视点绘制技术仍不能达到令人满意的效果,绘制的虚拟视点中大都会存在一些空洞和伪影^[6],影响最终的观看效果。目前,虚拟视点质量仍然采用传统的峰值信噪比(peak signal to noise ratio, PSNR)等方法来评价,这些方法没有抓住虚拟视点绘制失真的特点,不能真正反映 FVV 系统中进入用户眼中的虚拟视点图像对的感知质量。符合人眼主观感受的虚拟视点图像对的质量评价可作为反馈指导虚拟视点的绘制过程,进一步提升 FVV 系统中用户的感知质量。

传统立体图像评价方法多数采用 2D 图像的评价方法,例如 PSNR 和结构相似度指标(structural similarity index, SSIM),不能反映立体图像的双目特性,故不能很好地拟合人眼主观感受^[7,8]。Benoit 等人^[9]提出了一种立体图像质量评价算法,该方法利用传统的基于 SSIM 的方法获得左右图像的质量分数,将左右图像质量的平均值和视差图的质量分数结合起来得到最终的图像质量分数。Chen 等人^[10]从参考立体图像对和失真立体图像对中产生两个独眼图,然后分别计算两个独眼图的多尺度 SSIM 质量分数,将两个分数进行平均得到整体图像质量分数。Shao 等人^[11]根据立体图像特有的双目特性,对图像的不同区域分别求

取质量分数,利用加权求和的方式得到最终的质量分数。Zhou 和 Yu 等人^[12]提出一种无参考评价模型,基于人眼视觉感知特性提取预测特征,用 K -近邻的方法对相应的特征进行加权得出最终质量分数。上述立体图像的质量评价算法考虑的失真类型多数为 JPEG 压缩失真、JPEG2000 压缩失真、白噪声失真和高斯模糊失真等,这些失真往往均匀地分布在图像中。由于虚拟视点图像的失真多数为几何失真且集中在部分区域^[13],传统的立体图像质量评价算法不能够有效地评价 FVV 系统虚拟视点图像对的质量。

Battisti 等人^[14]提出一种虚拟视点质量评价算法,将图像分为互不重叠的图像块,对失真图像和参考图像匹配块作小波变换,通过比较两图像小波子带的统计特征来衡量图像的失真。Bosc 等人^[8]在分析了合成视点产生的失真类型和分布情况后,提出了一种针对虚拟视点合成的立体图像的质量评价方法,通过失真图像和原始图像之间的比较提取出失真严重区域并且只对该区域进行 SSIM 的计算,但其没有考虑双目特性在立体图像的质量评价中的影响。Jung 等人^[15]将双目特性因子融入虚拟视点评价算法中,提取左右失真图像失真严重区域,仅对该部分计算 SSIM 分数,将左右图像的分数进行平均作为最终立体图像的总质量,Jung 算法中忽略了深度感知因子对观看立体图像时的影响。

本文结合人眼视觉特性和虚拟视点合成立体图像特点提出一种自由视点视频系统中立体图像的质量评价算法。将参考图像和失真图像分别进行小波分解后再用对比敏感度函数(contrast sensitivity function, CSF)加权,将得到的两幅图像计算平均 SSIM 值;在失真图像的匹配区域中找出该图像视觉中心凹区域,在该区域上与参考图像相应区域计算平均相位相似度;在匹配区域上利用阈值提取出人眼能感知的失真区域,在该区域上与利用视差映射过来的另一视点图像相对应



区域计算平均 SSIM 值。用图像自身携带的信息量对左右图像基于上述过程分别得到的 3 部分质量分数对应加权求和得到立体图像的质量分数，最后将这 3 部分质量分数加权作为最终的质量分数。实验结果表明，本文算法对虚拟视点质量评价具有较好的效果。

2 自由视点视频系统中立体图像的质量评价算法

2.1 算法整体框架

本文提出的自由视点视频系统中立体图像的质量评价算法整体框架如图 1 所示。根据人眼视觉特性，分别计算左右视点虚拟视点图像基于图像内容、基于深度感知和基于双目特性的 3 个质量分数，然后利用左右视点虚拟图像的信息量进行加权得到立体图像的质量分数，最后再

将这 3 部分质量分数线性加权得最终的立体图像质量分数。图 1 中 Q_{1L} 、 Q_{2L} 、 Q_{3L} 、 Q_{1R} 、 Q_{2R} 、 Q_{3R} 分别为左图像基于内容的质量分数、左图像基于深度感知的质量分数、左图像基于双目特性的质量分数、右图像基于内容的质量分数、右图像基于深度感知的质量分数和右图像基于双目特性的质量分数， Q_1 、 Q_2 、 Q_3 分别为立体图像基于内容的质量分数、立体图像基于深度感知的质量分数和立体图像基于双目特性的质量分数， Q 为最终立体图像质量分数。

本文算法中，左右虚拟视点图像的质量求取方法一样，图 2 为左视点图像的算法框架，将参考左图像和失真左图像分别进行小波分解，利用 CSF 的带通特性对分解后的小波系数进行加权，用平均 SSIM 计算两者之间的相似性，得到 Q_{1L} ；找到匹配区域上视觉中心凹区域，在该区域上计

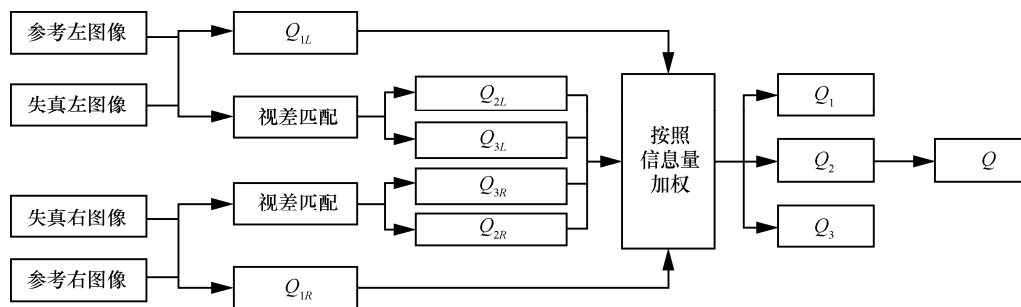


图 1 算法整体框架

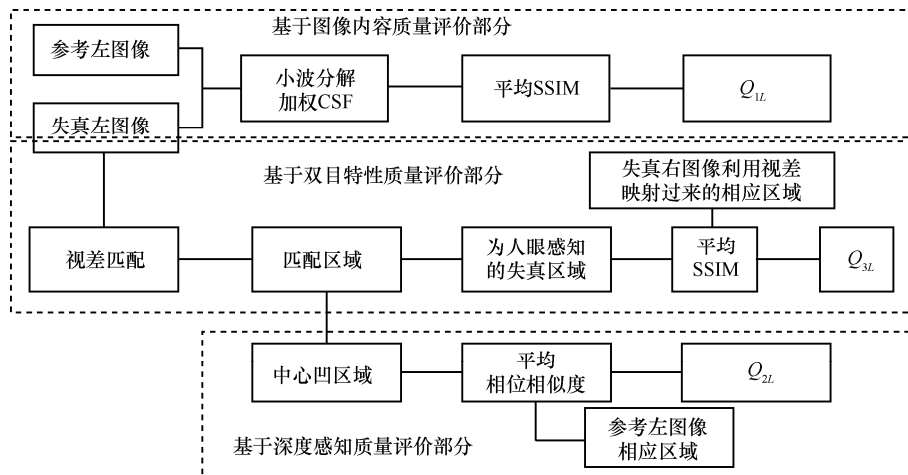


图 2 左视点图像算法框架

算参考左图像与失真左图像的平均相位相似度, 得到 Q_{2L} ; 利用阈值找出匹配区域上失真能为人眼所感知的区域, 与利用视差映射过来的失真右图像的相应区域计算平均 SSIM 值, 得到 Q_{3L} 。

2.2 基于内容的评价

由于人类视觉系统 (human visual system, HVS) 具有多通道特性、对比敏感度特性、掩蔽效应和视觉非线性定律等特性, 将符合人眼感知特性的影响因子加入评价体系中会得到不错的效果^[16]。人眼在观看一幅图像时, 会将图像分解为一组具有方向性的带通滤波器, 每个滤波器只对其中心频率附近一定区域的空间频率和方向做出响应, 小波分解后的数据具有此特性, 并且 CSF 曲线也具有带通特性, 只对一定的频率范围最敏感。据此, 本文对参考图像和失真图像分别进行小波分解, 然后再利用 CSF 对分解后不同空间频带的小波系数进行加权。最后用平均 SSIM 来计算失真图像与参考图像的相似度, 获得基于内容的质量分数。

令 X 与 Y 分别代表参考图像和失真图像, μ_X 与 μ_Y 分别代表参考图像与失真图像的均值, $l(X, Y)$ 为亮度对比函数, 计算方法为:

$$l(X, Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1} \quad (1)$$

σ_X 和 σ_Y 分别代表参考图像和失真图像的标准差, $c(X, Y)$ 为对比度函数, 计算方法为:

$$c(X, Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \quad (2)$$

σ_{XY} 为参考图像和失真图像的协方差, $s(X, Y)$ 为结构对比函数, 计算方法为:

$$s(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3} \quad (3)$$

其中, 式 (1)、式 (2)、式 (3) 中的 C_1 、 C_2 、 C_3 均为正常数。

SSIM (X, Y) 为结构相似度函数, β 、 α 、 γ 分别为 $l(X, Y)$ 、 $c(X, Y)$ 、 $s(X, Y)$ 的系数, 计算方法为:

$$\text{SSIM}(X, Y) = l(X, Y)^\alpha c(X, Y)^\beta s(X, Y)^\gamma \quad (4)$$

将参考图像和失真图像分别分成互不重叠的相同大小的图像块, 利用式 (4) 计算对应块的 SSIM, 再利用式 (5) 求得平均 SSIM 值, 即基于内容的质量分数:

$$Q_{1s}(X, Y) = \frac{1}{N} \sum_{i=1}^N \text{SSIM}(m_i, n_i), s = L, R \quad (5)$$

其中, m_i 与 n_i 分别代表参考图像和失真图像的各个图像块, N 为图像所包含的图像块个数。

2.3 基于深度感知的评价

在观看立体图像时, 除了画面本身质量会影响观看感受, 还有一点就是画面所提供的深度感, 这也是立体图像区别于 2D 图像的主要一点。在观看立体图像时, 人眼对于不同的区域会有不同的分辨率, 具有极高分辨率的一小块中心区域称为“中心凹区域”, 具有较差分辨率的大部分主要视觉区域称为“周围视觉”^[17]。根据参考文献[10-18], 可知参考图像和失真图像之间的相位特性不仅能有效地估计出图像的失真程度, 而且能反映人眼的深度感知。因此, 首先利用双目匹配区域找到人眼的“中心凹区域”, 只对该区域计算参考图像与失真图像的平均相位相似度, 以此来获得基于深度感知的质量分数:

$$Q_{2s}(X, Y) = \begin{cases} \frac{1}{K} \sum_{i=1}^K \frac{2ph_1(x_i, y_i) \cdot ph_2(x_i, y_i) + b_1}{ph_1^2(x_i, y_i) + ph_2^2(x_i, y_i) + b_1}, & \text{中心凹区域} \\ 0, & \text{其他} \end{cases}, s = L, R \quad (6)$$

其中, X 与 Y 分别代表参考图像和失真图像, $ph_1(x_i, y_i)$ 和 $ph_2(x_i, y_i)$ 分别代表参考图像相应区域和失真图像中心凹区域各像素点的相位, K 是中心凹区域所包含的像素点个数, b_1 是一个正常数。

2.4 基于双目特性部分

通过 Jung 等人^[15]的工作可知人眼在观看合成立体图像时, 对左右图像匹配区域的失真情况较敏感, 通过阈值的设定将匹配区域中人眼能感知



到的失真部分提取出来。

$$\text{Th}(x, y) = \frac{\ln[D(x, y) \cdot E(x, y) + 1]}{T_1} + \text{BJND}(x, y) \quad (7)$$

式(7)为阈值的选取,分为两部分,前半部分中 $D(x, y)$ 为相应图像的视差图, $E(x, y)$ 为利用视差图求得的左右失真图像的主要差异部分, T_1 为一正常数。后半部分中的 $\text{BJND}(x, y)$ 为双目恰可察觉模型,用于视差很小可以近似忽略不计情况下的阈值求取。图像中失真大于该阈值的部分即人眼能感知失真区域。

在立体图像对中只有单视点图像的失真不小于 $\text{BJND}(x, y)$ 时,人眼才能察觉到该立体图像对的失真。Jung 等人参考的 BJND 模型^[19]中考虑了亮度、对比度和加入均匀随机噪声 3 个影响因素,在左右图像给定的条件下,给出右图像的 BJND 为:

$$\begin{aligned} \text{BJND}_r(bg_L(I+d), eh_L(I+d), A_L(I+d)) \quad (8) \\ = A_{C, \text{limit}}(bg_L(I+d), eh_L(I+d)) \times \\ \left(1 - \left(\frac{A_L(I+d)}{A_{C, \text{limit}}(bg_L(I+d), eh_L(I+d))} \right)^\lambda \right)^{\frac{1}{\lambda}} \end{aligned}$$

其中, d 表示左右图像之间的视差, $bg_L(I+d)$ 表示

区域 I 的平均亮度值, $eh_L(I+d)$ 代表区域 I 的边缘梯度大小, $A_L(I+d)$ 表示左图像区域 I 所加的最大可容忍的随机噪声幅度, $A_{C, \text{limit}}(bg_L(I+d)$ 和 $eh_L(I+d)$) 是在左图像所加随机噪声为零的情况下,右图像所能容忍的最大随机噪声幅度,参数 λ 取值范围为 1.0~1.5。左视点的 BJND 也用同样的方法获得。

但本文结合虚拟视点绘制图像自身的特点,不同于普通立体图像失真的情况,绘制失真大都分布在特定的区域(去遮挡区域),图像中不存在均匀分布的随机噪声,所以上述阈值后半部分所采取的 BJND 模型不适用于虚拟视点绘制的图像,故将上述阈值的确定进行了一些改动,在计算左右图像的 BJND 时,将最后的随机噪声影响因子去掉,只考虑亮度掩蔽效应对比度掩蔽效应因子,使得该模型更符合虚拟视点绘制图像的特点,也使得计算过程更为简化,具体的计算式为:

$$\text{BJND}'(bg(I+d), eh(I+d)) = A_{C, \text{limit}}(bg(I+d), eh(I+d)) \quad (9)$$

式(10)为阈值的第 1 部分,考虑的是图像中不同大小的视差给阈值带来的影响:

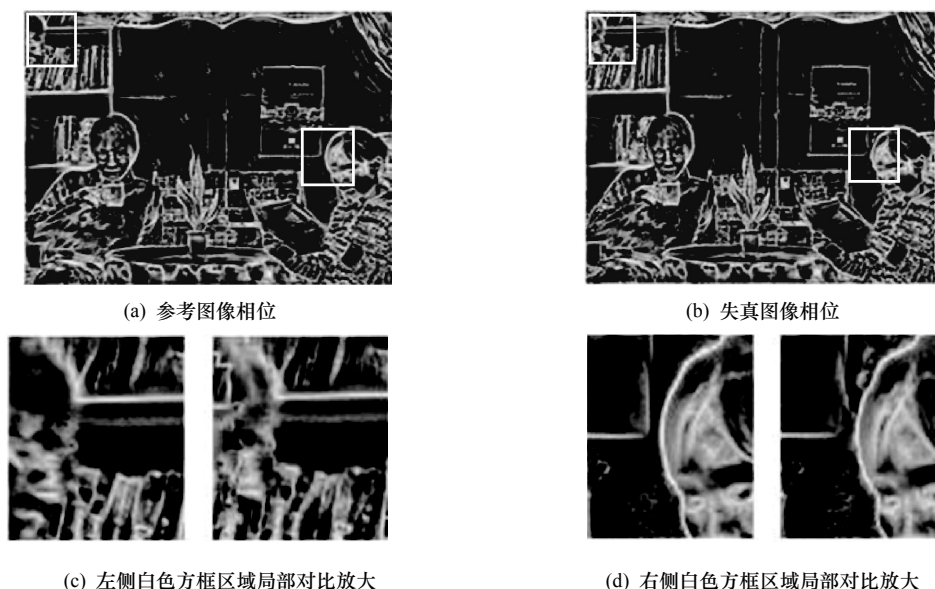


图3 参考图像与失真图像相位对比

$$Th_1(x, y) = \frac{\ln[D(x, y) \cdot E(x, y) + 1]}{T_1} \quad (10)$$

式(11)为阈值的第2部分,考虑的是当视差很小可以忽略的情况下阈值的选取:

$$Th_2(x, y) = A_{C, \text{limit}}(bg(I + d), eh(I + d)) \quad (11)$$

将上述两部分子阈值结合,得到最终的阈值 $Th(x, y)$:

$$Th(x, y) = Th_1(x, y) + Th_2(x, y) \quad (12)$$

运用上述阈值,提取出能为人眼所感知的失真区域,并只在该区域上与利用视差映射的另一视点图像的相应部分计算平均 SSIM 值,得到基于双目特性部分的质量分数:

$$Q_{3s}(F, G) = \begin{cases} \frac{1}{J} \sum_{i=1}^J SSIM(f_i, g_i), \text{人眼能感知的失真区域} & (s = L, R) \\ 0, & \text{其他} \end{cases} \quad (13)$$

其中, F 与 G 分别代表失真图像和另一视点失真图像, J 为失真图像中人眼能感知的失真区域所包含的像素点个数, f_i 代表失真图像中人眼能感知失真区域的各个像素点, g_i 代表另一视点失真图像利用视差得到的与失真图像中人眼能感知失真区

域相对应区域的各个像素点。图4为右图像根据上述阈值提取的人眼能感知失真区域的中间过程及其失真局部放大。

2.5 各部分权值确定

用 Logistic4 参数拟合客观质量评价结果,得到 $DMOS_p$ 值:

$$DMOS_p = \frac{a - b}{1 + \exp\left[-\frac{Q - c}{abs(d)}\right]} + b \quad (14)$$

其中, a 、 b 、 c 、 d 为常数。

根据参考文献[20],对于像 3D 虚拟视点图像对这样含有非对称失真的立体图像对来说,对左右图像赋相同权值的融合方法会带来一定的预测误差,且文中还提到利用信息量进行加权的方法,由于人眼对携带信息量多的部分较为关注,应对其赋予更高的权值。受此启发,图像熵能够反映其携带的信息量,可以通过计算输入图像的熵值分别给左右图像 3 部分质量分数赋以权重,得到立体图像的 3 部分质量分数,这样更加符合人眼相应的视觉特性。首先计算输入图像的直方图,然后按照熵的计算式进行计算:

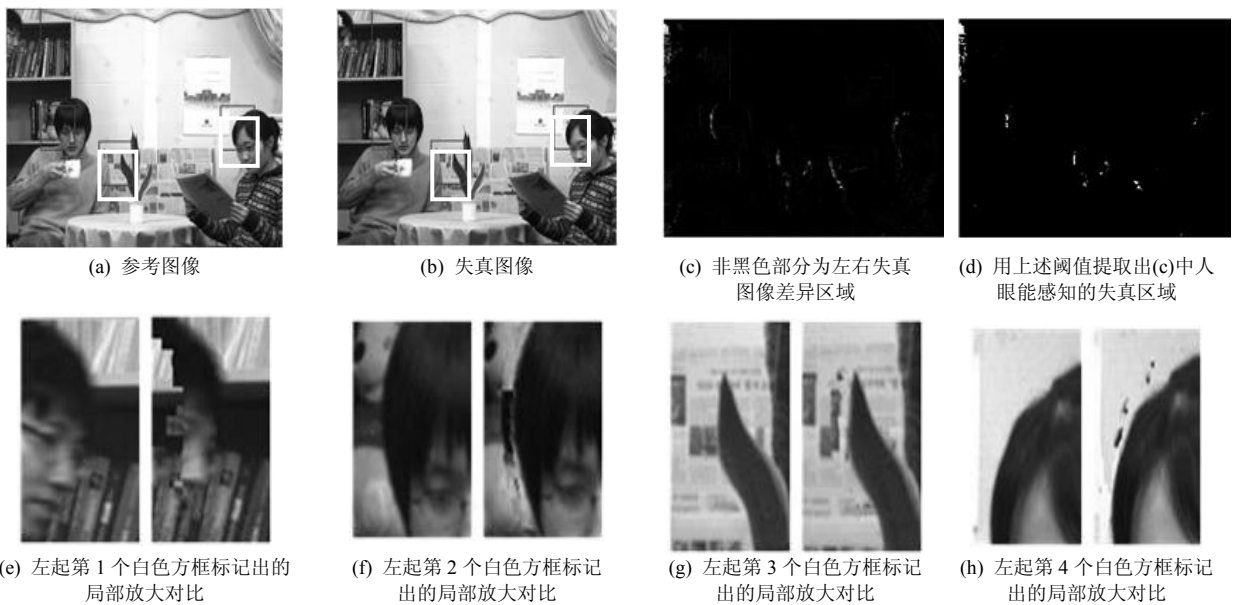


图4 失真图像中人眼能感知失真区域的提取过程及其失真局部放大



$$P_{uv} = \frac{h(u,v)}{M^2} \quad (15)$$

$$En = -\sum_{i=0}^{255} P_{uv} \lg P_{uv} \quad (16)$$

其中, $h(u,v)$ 表示反映空间特征的特征函数, u 表示像素的灰度值, v 表示邻域灰度均值, M 表示图像的尺度, En 表示所求图像的二维熵。图 5 为左右图像与其对应的熵。

从图 5 中可以看出, 由于左右图像存在一定的视差, 所包含的内容就有所差异, 且其所包含的失真为非对称失真, 以上原因使得左右图像熵图有所不同, 则所携带的信息量也就不同。根据上述信息加权理论, 在观看虚拟视点合成的立体图像时, 左右眼所占的权重因所含信息量的不同而不同, 故以图像熵值给左右图像的 3 部分质量分数分别赋相应权值。

左图像对应的质量分数所赋权值:

$$k_L = \frac{En_L^2}{En_L^2 + En_R^2} \quad (17)$$

右图像对应的质量分数所赋权值:

$$k_R = \frac{En_R^2}{En_L^2 + En_R^2} \quad (18)$$

则立体图像的各质量分数为:

$$Q_i = k_L Q_{iL} + k_R Q_{iR} \quad (i=1,2,3) \quad (19)$$

以线性加权的方式结合 Q_1 、 Q_2 和 Q_3 得到最终立体图像质量分数 Q :

$$Q = \omega_1 Q_1 + \omega_2 Q_2 + \omega_3 Q_3 \quad (20)$$

其中, ω_1 、 ω_2 和 ω_3 分别为 3 部分质量分数的权重值。

根据本算法所提出的 3 个影响因子, 给 ω_3 赋予一个初始值 ω_0 , 范围是[0~1], 调节步长设置为 0.01, 在考虑权值分配合理的情况下, 对 ω_1 和 ω_2 赋以不同的权重比例, 观测不同权值情况下 Spearman 相关系数 (spearman rank order correlation coefficient, SROCC) 评价指标的变化, 在其输出为优的范围内综合考虑其他性能指标, 将最合理的那一组权重值分别赋给 Q_1 、 Q_2 和 Q_3 , 得到 Q 。如图 6 所示, 横坐标是 ω_3 的不同取值, 纵坐标是 SROCC 指标的输, 在 ω_1 与 ω_2 的赋值比例设为 3:1, ω_3 权值设定为 0.6 左右时, SROCC 指标为优输出范围, 综合考虑后, 最终将 ω_1 与 ω_2 的比例设为 3:1, ω_3 取权值为 0.61。



(a) 左图像



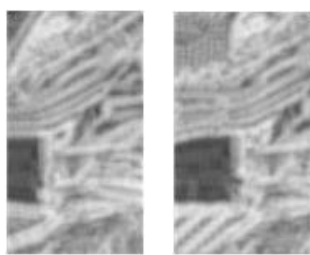
(b) 右图像



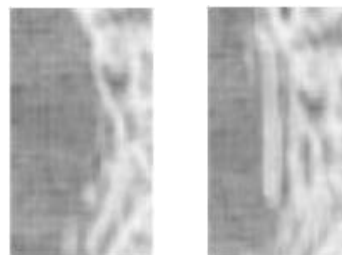
(c) 左图像熵



(d) 右图像熵



(e) 左起第 1 个方框标记出的局部放大对比



(f) 左起第 2 个方框标记出的局部放大对比

图 5 左右图像及其对应的熵

3 实验结果与讨论

3.1 数据库

实验采用 Jung^[21]提供的虚拟视点立体图片数据库。该数据库中包含7个不同的图片组,其中4个图片组是来自 Middlebury^[21,22]数据库的 Aloe、Dolls、Reindeer 和 Laundry,剩下的3组是来自 MPEG 3DV 多视点加深度序列的 Lovebird1 (145 帧)、Newspaper (41 帧)、Bookarrival (85 帧),每组图片包含用4种不同绘制算法(Ahn、Criminisi、Inter-view Consistent、VSRS)获得的不同失真类型和程度的4个不同视点的失真图片。这些失真图片相邻视点间两两组合,构成了72对失真的虚拟视点立体图像对。同时,该库还提供了每个立体图像对的主观打分值,如图6所示。

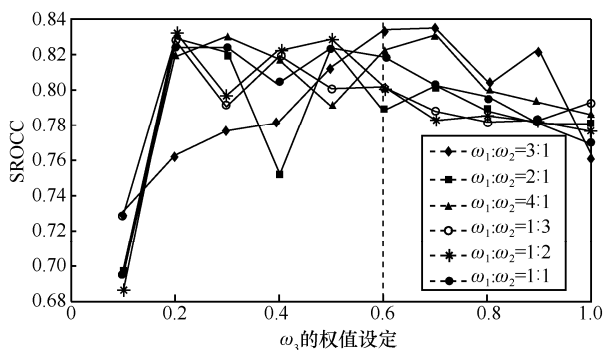


图6 不同权值设定下的 SROCC 指标变化情况

3.2 客观算法的性能测试

本文采用 SROCC、Pearson 相关系数 (pearson linear correlation coefficient, PLCC) 以及均方根误差 (root mean squared error, RMSE) 3 种评价指标衡量本文所提算法性能。实验结果见表 1。

表 1 给出了 8 种方法在 Jung 提供的数据库上的各项性能指标,从实验结果上看,本算法基于人眼感知提取出的 3 个影响因子与人眼主观感受拟合性较好,PLCC 和 SROCC 两个指标均优于其他算法,但 RMSE 指标有所升高,为进一步探究本算法的有效性,图 7 给出了部分图片组主观打分值与客观评价值的散点图,结合散点图对上述

表 1 不同客观评价算法的 3 种指标对比

评价指标		PLCC	SROCC	RMSE
客观评价算法	SSIM	0.351	0.381	13.585
	MS_SSIM	0.637	0.647	11.180
	VIF	0.388	0.299	13.373
	Benoit	0.452	0.438	12.944
	3DSwIM	0.527	0.499	12.330
	Bosc	0.642	0.596	11.135
	Chen	0.687	0.695	10.631
	Jung	0.826	0.829	8.181
	本文算法	0.855	0.834	11.28

实验结果做以下分析。

传统 2D 评价算法并未将立体图像特有的人眼感知特性加入评价算法中,所以算法与人眼主观感受拟合性不好;普通两视点立体图像评价算法不能够反映虚拟视点绘制立体图像失真的特殊性,算法与人眼主观感受拟合性也不好;Battisti、Bosc 和 Jung 在虚拟视点绘制立体图像的基础上提出客观评价算法,Battisti 和 Bosc 未将双目特性考虑在内,而 Jung 忽略了深度感知的影响。与以上算法相比,本文较充分考虑到人眼视觉特性。

本算法所用的数据库中每组图片提供的相应视点的深度图都存在一定的失真,则利用深度与视差的关系转化来的视差图就存在相应失真。尤其是 Dolls 和 Laundry 这两组图片,其视差图存在误差的多数区域刚好是待评价图像的失真区域,所以在利用视差大小设定阈值求取失真为人眼感知的区域(失真严重区域)就会存在较大的偏差。同时,视差的失真也会造成提取的左右失真图像匹配区域的不准确,进而导致深度感知因子存在偏差。故从图 7(a)和图 7(b)中可以看到 Dolls 和 Laundry 这两组图片相对于其他图片组拟合性稍差了些。

从图 7(c)可以看出,算法对 Lovebird1 (145 帧)

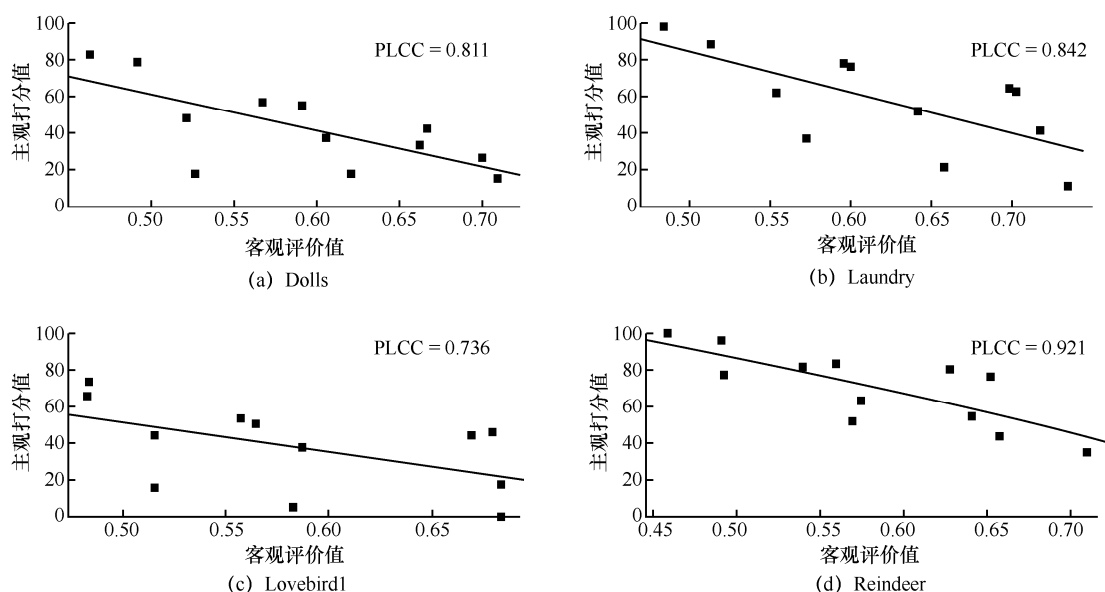


图7 部分图片组主观打分值与客观评价值的散点

图片组效果不是很理想, 因其图片背景的某些临近区域亮度存在明显差异, 而本算法相对于其他算法对亮度分量更加敏感, 因提取的 3 个影响因子最终计算质量分数时所用的结构相似度和相位相似度以及多个中间过程均与图像的亮度分量有关, 这种亮度的不准确降低了求取客观质量分数的准确性。

4 结束语

本文从人眼视觉特性出发, 基于 3 部分评价自由视点视频系统中立体图像的质量, 分别是基于图像内容、基于深度感知和基于双目特性, 较充分地考虑了人眼观看立体图像时的影响因子, 并根据图像本身所携带的信息量赋予不同视点相应的权重, 能够很好地模拟人眼观看立体图像时的情景。上述实验结果分析中提到的原因造成了整体 RMSE 指标的升高, 其也是 PLCC 和 SROCC 两个指标虽有所升高, 但结果仍有待提高的原因之一, 从侧面反映出本文所提出的自由视点视频系统中立体图像的质量评价算法在某种程度上具有一定的不可靠性, 应进一步优化本算法, 获得更强的顽健性, 将是下一步研究工作的方向。

参考文献:

- [1] 陈晓贝, 魏克军. 全球 5G 研究动态和标准进展[J]. 电信科学, 2015, 31(5): 16-19.
CHEN X B, WEI K J. Global research and standardization progress of 5G[J]. Telecommunications Science, 2015, 31(5): 16-19.
- [2] 卢卫, 陆希玉. 4G 时代移动互联网的发展趋势[J]. 电信科学, 2014, 30(5): 51-54.
LU W, LU X Y. Mobile internet trends in 4G era[J]. Telecommunications Science, 2014, 30(5): 51-54.
- [3] 林金桐, 许晓东. 第五代移动互联网[J]. 电信科学, 2015, 31(5): 7-14.
LIN J T, XU X D. The 5th generation of mobile internet[J]. Telecommunications Science, 2015, 31(5): 7-14.
- [4] COLLET A, CHUANG M, SWEENEY P, et al. High quality streamable free viewpoint video[J]. ACM Transactions on Graphics, 2015, 34(4): 1-13.
- [5] FEHN C. Depth image based rendering (DIBR), compression and transmission for a new approach on 3D TV[J]. Proceedings of the 2004 SPIE Stereoscopic Displays and Virtual Reality Systems XI, 2004(5291): 93-104.
- [6] CHO J H, SONG W, CHOI H, et al. Hole filling method for depth image based rendering based on boundary decision[J]. IEEE Signal Processing Letters, 2017, 24(3): 329-333.
- [7] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004(13): 600.
- [8] BOSCH E. Towards a new quality metric for 3D synthesized view assessment[J]. IEEE Journal of Selected Topics in Signal Pro-

- cessing, 2011, 5(7): 1332-1343.
- [9] BENOIT P, CALLETL, CAMPISI P, et al. Using disparity for quality assessment of stereoscopic images[C]//2008 IEEE International Conference Image Processing, October 12-15, 2008, San Diego, California America. Piscataway: IEEE Press, 2008: 389-392.
- [10] CHEN M J, SU C C, KWON D K, et al. Full reference quality assessment of stereopairs accounting for rivalry[J]. Signal Processing Image Communication, 2013, 28(9): 1143-1155.
- [11] SHAO F, LIN W, GU S, et al. Perceptual full reference quality assessment of stereoscopic images by considering binocular visual characteristics[J]. IEEE Transactions on Image Processing, 2013, 22(5): 1940-1953.
- [12] ZHOU W, YU L. Binocular responses for no reference 3D image quality assessment[J]. IEEE Transactions on Multimedia, 2016, 18(6): 1077-1084.
- [13] CHEN F, JIAO R, PENG Z, et al. Virtual view quality assessment based on shift compensation and visual masking effect[J]. Journal of Visual Communication & Image Representation, 2017(43): 41-49.
- [14] BATTISTI F, BOSCH E, CARLI M, et al. Objective image quality assessment of 3D synthesized views[J]. Signal Processing Image Communication, 2015, 30(C): 78-88.
- [15] YONG J J, KIM H G, YONG M R, et al. Critical binocular asymmetry measure for the perceptual quality assessment of synthesized stereo 3D Images in view synthesis[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2016, 26(7): 1201-1214.
- [16] GAO X, LU W, TAO D, et al. Image quality assessment and human visual system[J]. Proceedings of The International Society for Optical Engineering, 2010, 7744(5): 17-19.
- [17] RIBEIRO F, FLORÊNCIO D. Region of interest determination using human computation[C]//IEEE International Workshop on Multimedia Signal Processing, October 17-19, 2011, Hangzhou, China. Piscataway: IEEE Press, 2011: 1-5.
- [18] ZHANG L, MOU X. FSIM: a feature similarity index for image quality assessment[J]. IEEE Transactions on Image Processing, 2011, 20(8): 2378-2386.
- [19] ZHAO Y, CHEN Z, ZHU C, et al. Binocular just noticeable difference model for stereoscopic images[J]. IEEE Signal Processing Letters, 2011, 18(1): 19-22.

- [20] WANG J, REHMAN A, ZENG K, et al. Quality prediction of asymmetrically distorted stereoscopic 3D images[J]. IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society, 2015, 24(11): 3400-3414.
- [21] IVY LAB Stereoscopic 3D image database for quality assessment of synthesized stereo images[EB/OL]. (2014-10-01)[2017-06-03]. <http://ivylib.kaist.ac.kr/demo/3DCBA/3DCBA.htm>.
- [22] Middlebury stereo datasets[EB/OL]. (2015-06-29)[2017-06-03]. <http://vision.middlebury.edu/stereo/data/>.

[作者简介]



王士培 (1994-), 女, 宁波大学信息科学与工程学院硕士生, 主要研究方向为虚拟视点绘制及其质量评价。



彭宗举 (1973-), 男, 博士, 宁波大学信息科学与工程学院教授, 主要研究方向为三维视频信号处理与编码。



陈芬 (1973-), 女, 博士, 宁波大学信息科学与工程学院副教授, 主要研究方向为三维视频信号处理与编码。

蒋刚毅 (1964-), 男, 博士, 宁波大学信息科学与工程学院教授, 主要研究方向为三维视频信号处理与编码。

郁梅 (1968-), 女, 博士, 宁波大学信息科学与工程学院教授, 主要研究方向为三维视频信号处理与编码。